Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Recombinant cloning strategies for protein expression

Patrick HN Celie[1], Annabel HA Parret[2] and Anastassis Perrakis[1]

A variety of methods to create specific constructs for protein expression, in a broad range of organisms, are available nowadays. Restriction enzyme-free, ligation-independent and recombinase-based cloning methods have enabled high-throughput protein expression for structural and functional studies. These methods are also instrumental for modification of target genes including gene truncations, site-specific mutagenesis and domain swapping. Here, we describe the most common cloning techniques that are currently at hand for recombinant protein expression studies, including a brief overview of techniques associated with co-expression experiments. We also provide an inventory of many of the available reagents for the various cloning methods, and an overview for some computational tools that can help with the design of expression constructs.

Addresses
[1] Division of Biochemistry, Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands
[2] European Molecular Biology Laboratory, Hamburg Unit, Notkestrasse 85, 22607 Hamburg, Germany

Corresponding author: Celie, Patrick HN (p.celie@nki.nl)

Current Opinion in Structural Biology 2016, **38**:145–154

This review comes from a themed issue on **New Constructs and expression of proteins**

Edited by **Rob Meijers** and **Anastassis Perrakis**

http://dx.doi.org/10.1016/j.sbi.2016.06.010

Introduction

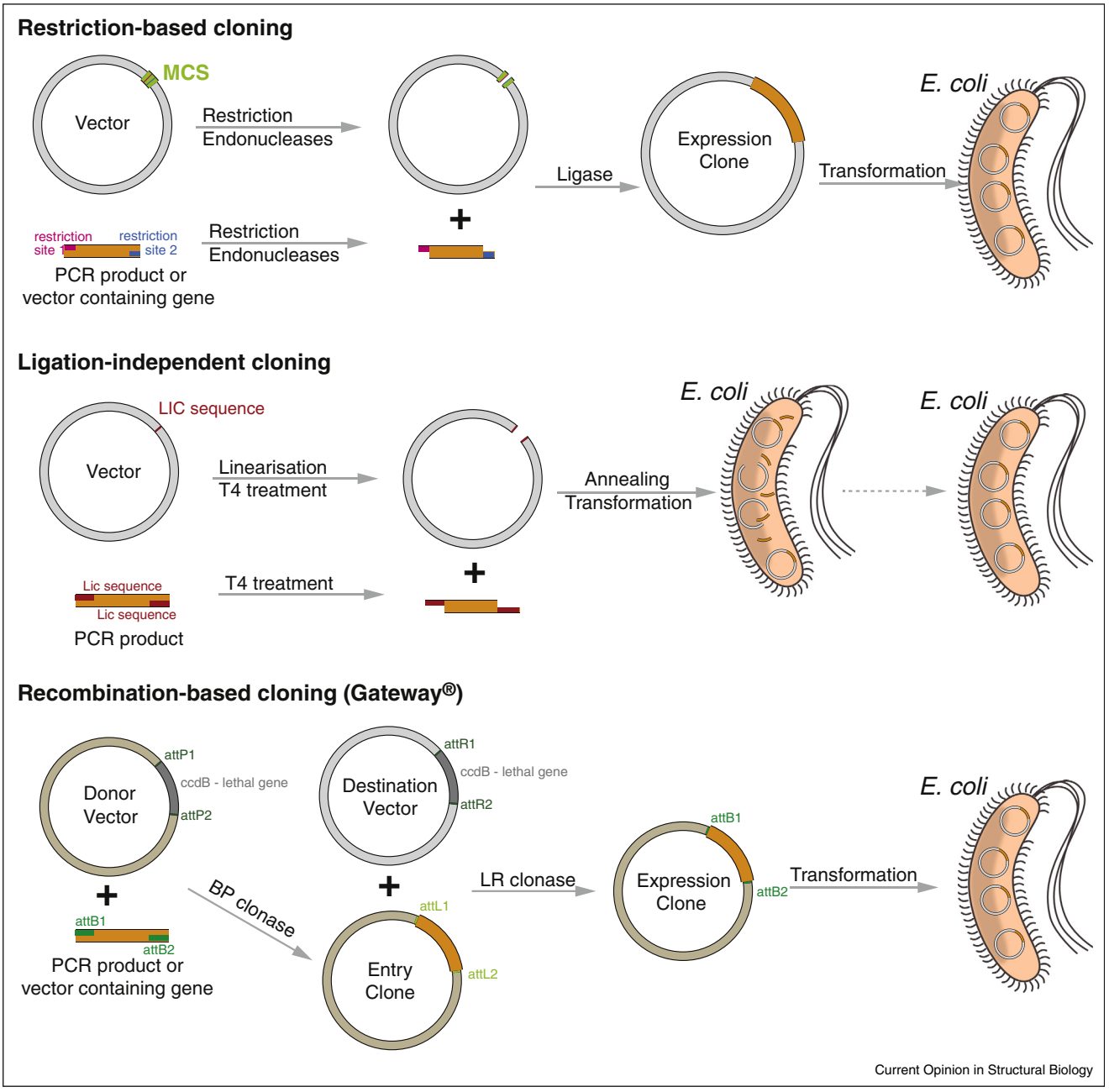
Molecular cloning techniques have advanced dramatically since the discovery of the first restriction endonucleases. Recombinant DNA technology is nowadays considered a routine practice. DNA isolation and amplification, Polymerase Chain Reaction (PCR), molecular cloning, and — more recently — genome editing have become standard procedures.

In the early 1970s, discovery of the restriction enzymes HindIII [1] and EcoRI [2] led to a significant breakthrough in the development of recombinant DNA technology and pioneered the first cloning experiments transferring DNA fragments from one bacterial strain to another, using a carrier plasmid [3]. Cloning of a complete bacterial gene followed [4], and the discovery that DNA from different species can be cloned and propagated in another species [5] signified the beginning of the heterologous recombinant gene expression technologies. In the 1980s, the breakthrough invention of PCR to amplify DNA fragments *in vitro* [6], allowed advanced cloning methodologies to emerge. Routine synthesis of large DNA fragments of specific sequence [7] that became mainstream and affordable in the last decade, was another landmark in laboratory practice. In this review, the most popular cloning techniques that are of direct relevance to protein expression will be presented and discussed.

Restriction and ligation-based cloning

At present, the restriction enzyme repertoire consists of more than 200 highly specific and commercially available proteins. Likewise, many diverse cloning vectors exist that enable ligation of DNA fragments digested with specific restriction enzymes into their multiple cloning site (MCS), consisting of unique restriction sites (Figure 1). However, cloning can become complicated when genes contain internal restriction sites that are also present in the MCS. In addition, when multiple expression constructs have to be made, for example, for protein production in various host organisms or when different affinity and solubilizing tags need to be tested, compatibility between the MCS of different vectors is an issue. In addition, many restriction enzymes often do not cleave linear PCR products well, but prefer circular DNA. Finally, 'seamless' cloning is not always possible. The selected restriction site(s) may introduce additional nucleotides into the coding sequence, causing a scar in the gene which leads to incorporation of non-native amino acid residues into the expressed protein. Various solutions have been invented to tackle some of these cloning issues. Multiple-host vectors have been designed to allow protein expression from the same construct in *E. coli*, insect, and mammalian cells (pTriEX™ vectors; EMD Millipore). To enhance cleavage specificity of DNA fragments, a deoxyinosine nucleotide can be incorporated in flanking DNA sequences, to allow cleavage by Endonuclease V [8]. Type II restriction endonucleases, which cleave outside their recognition sequence markedly facilitate seamless cloning [9]. This approach has been adopted by Golden Gate cloning [10–12] and is commercialized by NEB. TOPO®-TA (sub) cloning (Thermo Fisher Scientific) is a restriction-independent cloning method. It takes advantage of the specific feature of the Taq polymerase, which adds a 3′ overhanging adenine (A) base to a double-stranded DNA fragment. This allows efficient topoisomerase-assisted ligation of a

Figure 1



Schematic representation of three different cloning techniques. Restriction-based cloning requires restriction endonucleases to create compatible overhangs in vector and insert which can be joined by a ligase. In ligation-independent cloning (LIC) [14], compatible overhangs in insert and vector are created by exonuclease activity of T4 polymerase and the ligation of the fragments occurs inside *E. coli* upon transfection of the fragments. Recombination of two DNA fragments, like in the Gateway® system [26,27,42], requires specific sequences (*att*) within insert and vector, which are recognized by the recombinase enzyme mixtures (Clonase). Within the Gateway® system, typically a subset of entry clones is created which can be used for subsequent cloning of the insert into destination (expression) vectors.

PCR fragment into a linearized vector comprising a thymidine (T) overhang at the 5′ end. Directionality issues, as well as the high cost of vectors covalently linked to topoisomerase, have compromised the popularity of this solution; it has often been used, however, to transfer a PCR product to an 'intermediate' vector, avoiding issues with restriction enzymes being inefficient in cutting PCR products.

Notably, the USER cloning™ strategy from NEB [13] is unidirectional, although it requires a specific enzyme mixture, and DNA fragments have to be amplified with